Data-independent proteomic screen identifies novel tamoxifen agonist that mediates drug resistance

Published in final edited form as:
J Proteome Res. 2011 October 7; 10(10): 4567–4578. doi:10.1021/pr2004117.

Data-independent Proteomic Screen Identifies Novel Tamoxifen Agonist that Mediates Drug Resistance

Shawna Mae Hengel[1,*], **Euan Murray**[2], **Simon Langdon**[2], **Larry Hayward**[2], **Jean O'Donoghue**[2], **Alexandre Panchaud**[1], **Ted Hupp**[2], and **David R. Goodlett**[1,**]

[1]University of Washington, Department of Medicinal Chemistry, Box 357610, Seattle, WA 98195-7610, USA

[2]University of Edinburgh, Crewe Road South, Edinburgh, EH4 2XR, United Kingdom

Abstract

A label-free quantitative variation of the recently developed data-independent shotgun proteomic method **P**recursor **A**cquisition **I**ndependent **F**rom **I**on **C**ount (PAcIFIC) was used to identify novel proteins implicated in cancer progression and resistance. Specifically, this screen identified the pro-metastatic protein anterior gradient 2 (AGR2) as significantly up-regulated in tamoxifen treated cells. Highlighting the need for direct proteome profiling methods like PAcIFIC, neither data-dependent gas-phase fractionation nor a transcriptomic screen detected AGR2 protein/transcript at significantly up-regulated levels. Further cell-based experiments using human cancer cell lines and *in vivo* xenografts confirmed the PAcIFIC hypothesis that AGR2 is up-regulated in MCF-7 cells post tamoxifen treatment, and that it is implicated in drug resistance mediation.

Keywords

tamoxifen; anterior gradient 2 homolog; AGR2; PAcIFIC; data independent

Introduction

Cancer progression is driven in part by progressive accumulation of mutations in selected proto-oncogenes and tumor suppressor genes that drive clonal cell evolution. Recent cancer genomic sequencing has identified thousands of mutations in the cancer genome, some of which have, or can, promote cancer cell survival.[1–4] The cancer genome sequence itself could be considered "static" and although it forms a record of the mutations that occurred during the cancer progression sequence, this record may not reflect *a priori* the most probable drug targets. Dynamic changes in the cancer cell transcriptome and proteome may better reflect pathways or targets of interest on which to focus anti-cancer drug development. Transcriptome approaches have been successful in identifying a large range of genes up-regulated during metastasis and angiogenesis and continue to provide new insights into clinically relevant carcinogenic mechanisms and drug targets. However, despite these successes using transcriptomics to identify novel resistance factors, there can be up to 80%

[**]Requests for reprints: David R. Goodlett Ph. D. University of Washington, Box 357610, Seattle, WA 98105, goodlett@u.washington.edu Phone: 206-616-4586 Fax: 206 685-3252.
[*]Current Address Pacific Northwest National Laboratory P.O. Box 999 MSIN K8-98, Richland, WA 99352 USA

Conflict of Interest

There are no financial or commercial conflicts of interest for any of the authors that contributed to this body of work.

Supporting Information

Supporting Information Available: This material is available free of charge via the Internet at http://pubs.acs.org

disagreement in transcript and protein abundance.[5–6] This discordance has dramatically increased demand for proteomic screens to determine the active cellular processes downstream from transcriptional events.

The many variations on the theme of shotgun proteomics used for proteome screens all utilize upfront fractionation procedures prior to protein sequencing via tandem mass spectrometry. These approaches include: **i)** SDS-PAGE, **ii)** chromatography, e.g. strong cation exchange[7] to enrich for broad classes of proteins and/or peptides and **iii)** affinity matrices that can enrich or deplete protein-specific chemical moieties like phosphate or high abundance proteins.[8] These methods typically produce many fractions, each of which is then subsequently analysed by data-dependent acquisition (DDA) driven HPLC-MS methods which are known to under sample the peptides present in each fraction.[9] Thus, no matter how much fractionation is carried out, the end result is under sampling of each of the fractions by DDA HPLC-MS/MS methods. Quantitative modifications of these approaches include *ex-vivo* or *in vivo* labelling of proteins using fluorescent or stable isotopic chemical tags in order to improve the sensitivity of the differential protein expression profile. The remaining difficulty in applying these standard fractionation techniques followed by proteomic profiling approaches to cancer biology and clinical medicine is the upfront manipulation involved in processing for electrophoresis, liquid or gas phase separations, and *in vivo* or *ex-vivo* labelling of proteins.

In addition to general proteomic profiling to define the components of a system, knowledge of the relative protein abundance between experimental conditions is crucial to understand the biological mechanisms that result from perturbations to the system. Unfortunately, many of the most popular quantitative proteomic screening methods preclude direct analysis due to requirement for fractionation (e.g. strong cation exchange separation of peptides) or incorporation of stable isotopes, both of which can result in loss of protein during preparation. The most popular methods include: **i) S**table **I**sotope **L**abeling with **A**mino acids in cell **C**ulture (SILAC)[10] which, due to the *in vivo* labelling protocol, is not possible using primary human clinical biopsies, and **ii)** a large number of variations on the theme of *ex vivo* labelling methods (e.g. iTRAQ[11–12] and TMT[13]) each of which requires fractionation prior to HPLC MS/MS analysis and concomitantly expands the number of samples that must be examined often complicating experimental design.Advances that streamline proteomic screens making them more similar to direct analysis without fractionation of transcriptomics would facilitate clinical proteomics.

Recently a novel, data-independent acquisition (DIA), tandem mass spectrometry based proteome screen referred to as **P**recursor **A**cquisition **I**ndependent **F**rom **I**on **C**ount (PAcIFIC)[14] was demonstrated to increase the identified proteomic dynamic range by an order of magnitude when compared to the standard DDA shotgun proteomic methods. The dynamic range benefit of PAcIFIC was gained by excluding pre-fractionation and use of direct systematic interrogation of all *m/z* “channels” for the presence of peptides regardless of the observation of a precursor ion. In combination this provides the DIA PAcIFIC method with two advantages over DDA methods because **1)** minimal sample handling decreases nonspecific losses of peptides from low copy number proteins and **2)** systematic interrogation of all *m/z* channels allows detection of peptides obscured by chemical noise. A key result of the initial PAcIFIC publication was to demonstrate that more information was available by thorough examination of a single serum sample rather than incomplete DDA examination of dozens of fractions. The ultimate goal of PAcIFIC analysis is to provide a simplified, automated, direct and systematic approach for proteomic screening that is more akin to genomic sequencing than other proteomic methods.

Here we describe results from the application of a quantitative PAcIFIC screen to a current problem in breast cancer treatment; tamoxifen resistant cancers. Breast cancer is the most commonly diagnosed cancer in women and it is estimated that approximately 10% of women will suffer from the disease at some point in their lifetime.[15] Improvements in breast cancer treatment over the last few decades have led to an overall decrease in mortality of the disease, with adjuvant anti-estrogen therapies (e.g. tamoxifen) effectively treating a majority of ER-positive breast cancer patients.[16] Despite these successes, approximately 50 percent of patients have intrinsic resistance to tamoxifen therapy and most of the remainder will acquire resistance to the drug, leading to metastatic cancer.[17–18] A further understanding of the tamoxifen mechanism of action and mechanisms of intrinsic, or acquired, resistance at the proteome level will aid in the development of modified disease treatment. As tamoxifen resistance and agonistic functions have been observed, we asked whether we could identify novel proteins up-regulated after tamoxifen treatment of MCF-7 breast cancer cells using the systematic PAcIFIC proteomic screen. Here we demonstrate the importance of increased dynamic range for identifying unique proteins by PAcIFIC[14] analysis that were not identified by DDA genome-based gas-phase fractionation (GPF),[9] and the incorporation of quantitative label-free spectral counting to identify these proteins of interest in tamoxifen resistance.

Experimental Procedures

Unless otherwise stated, all chemicals were purchased from Sigma. Whole genome transcriptome analysis using the Affymetrix platform with RNA from untreated and tamoxifen-treated cells was performed by Almac Diagnostics (Craigavon, UK). Doxorubicin and cisplatin were acquired from Calbiochem (Gibbstown, New Jersey).

Cell Culture/Xenografts/Target Validation

MCF-7 breast cancer cells (ECACC European Collection of Cell Cultures, Salisbury, Wiltshire (UK)) were cultured in DMEM supplemented with 10% FCS in a humidified incubator in 5% CO_2 at 37°C. Cells were plated at a density of 1.5×10^5 cells per well in six-well tissue culture plates 1 day prior to treatment. Tamoxifen was dissolved in DMSO to give a concentration of 1mM. For treatment of cells, this was diluted 1:1000 in DMEM 10% FCS to give a final tamoxifen concentration of 1μM. Cells were treated with 1μM tamoxifen or DMSO only (control) for a total of 96 hours, with the media and tamoxifen being replaced after 48 hours. Cells were harvested in ice cold PBS. Tamoxifen and control incubations were performed in parallel biological duplicates.

Cell survival or death using the AGR2-negative and positive cell panel with exposure to doxorubicin was determined using an MTT assay.[19] For xenograft studies, H1299 cell lines(American Type Culture Collection (ATCC), Manassas, VA), either anterior gradient homolog protein 2 (AGR2)-negative or transfected to acquire AGR2,[20] were processed in a xenograft system in the absence and presence of cisplatin, as previously described.[21–22] Briefly, transfected H1299 cells (5 × 106 cells / implant) were initially injected subcutaneously into the flanks of groups of adult female nude mice to establish the xenograft lines. Animals were at least 8 weeks old at the time of experimentation and were maintained in negative pressure isolators (La Calhene, UK). For experiments, tumor fragments were implanted subcutaneously into both flanks of nude mice and allowed to grow to 4–6 mm in diameter (over a period of approximately 1 month). Animals were then allocated to cisplatin treatment (5 mg/kg ; i.p.) or control (5 mice/group) groups and treatment was commenced (defined as day 0). Groups contained 8–10 tumors. Tumor size was measured twice weekly using calipers and the volume calculated according to the formula $\pi/6 \times \text{length} \times \text{width}^2$. Relative tumor volumes (%) were then calculated for each individual tumor by dividing the tumor volume on day t (V_t) by the tumor volume on day 0

(V_0) and multiplying by 100. Mean tumor volumes were then obtained. Ethical approval was granted by the University of Edinburgh ethical review committee and experiments were conducted according to UKCCCR guidelines under a Home Office license.[23]

For western blot analysis cells were lysed in the following buffer: 150mM NaCl, 50mM Tris HCl pH 8.0, 50mM NaF, 5mM EDTA, 1% NP-40 plus 1:100 phosphatase inhibitor cocktail and 1:100 protease inhibitor cocktail. AGR2 was detected using a polyclonal antibody to AGR2 raised in rabbit (Moravian Biotechnology Ltd., Brno Czech Republic).

For RT-PCR, total RNA from cells was extracted and purified using an RNeasy Mini kit (Qiagen, Hilden, Germany) with QIAshredder to homogenize lysate. cDNA was generated using the Omniscript RT kit (Qiagen, Hilden, Germany) and finally AGR2 and GAPDH were detected by PCR using HotStarTaq DNA polymerase (Qiagen, Hilden, Germany) with the following oligonucleotide primers; AGR2 forward: GCTCCTTGTGGCCCTCTCCTACAC, AGR2 reverse: ATCCTGGGGACATACTGGCCATCAG. GAPDH forward: GTCAGTGGTGGACCTGACCT, GAPDH reverse ACCTGGTGCTCAGTGTAGCC.

Mass spectrometry

After cells were lysed in 50 mM ammonium bicarbonate by alternating three cycles of sonication and cooling, insoluble fragments were removed by centrifugation. Protein concentration was determined with a standard BCA assay, and an excess of over 100 µg from each sample were allotted for digestion. Urea was added to a final concentration of 6M, and 7 uL of 1.5 M Tris pH 8.8 were added to maintain a basic solution during reduction in the 100 uL of solution. Proteins were reduced for 1 hr at 37° C with a final concentration of 5 mM tris(2-carboxyethyl)phosphine. Cysteine residues were alkylated by the addition of 20 uL of 200 mM iodoacetamide and incubation for 1 hr in the dark, followed by the addition of 20 uL of 200 mM dithiothreitol and one-hour incubation. Samples were diluted with 900 µL of 50 mM ammonium bicarbonate and sequencing-grade trypsin was added at a protein to enzyme ratio of 1:50. Digestion proceeded overnight at room temperature. Samples were desalted using a Vydac silica C_{18} macrospin column (The Nest Group, Southborough, MA).

Two types of HPLC-MS/MS data acquisition schemes were run in duplicate: 1) DDA genome-based GPF[9] and 2) a modified version of the DIA PAcIFIC method.[14]. All acquisitions were carried out using either the linear ion trap (LTQ) of a LTQ-Orbitrap hybrid mass spectrometer or a standalone LTQ mass spectrometer (Thermo Scientific Corp., San Jose, CA). Modifications to the published PAcIFIC analysis included increasing the mass range covered per method file from 15 to 22.5 *m/z* units. Consecutive scans remained 1.5 units apart with an isolation width of 2.5 *m/z* as previously described. Genome-based GPF was conducted using MS scan ranges of: 350–520, 515–690, 685–970, and 965–2000 *m/z*. For all acquisition methods, approximately 1 µg of peptides per injection was separated online using home-packed columns (0.75 µm i.d. × 11cm; 100Å Magic C18AQ (Michrom Bioresources, Auburn, CA)) and a Nano-Acquity (Waters, Milford, MA) or 1100 binary HPLC (Agilent, Santa Clara, CA) with a linear gradient of 5–35% ACN over 60 minutes. For each full PAcIFIC analysis 45 µg of total peptide digest was required per sample/replicate. Electrospray ionization voltage was applied via a gold wire and liquid microtee junction.[24] All ions were measured in positive mode and source parameters were optimized for each system using the manufacturer-suggested tune solution. Automatic Gain Control (AGC) was used to maintain constant ion populations at $3e^4$ and $1e^4$ for precursor ion scans (MS) and CID fragment ion scans, respectively.

Protein Identification/Quantification

Raw data files were converted to the mzXML format using ReAdW.exe version 4.0.2 (Institute for Systems Biology, Seattle, WA) with default parameters. Proteins were identified by SEQUEST,[25] version 27, specifying optional cysteine alkylation in the search parameter file. Other search parameters included: precursor mass tolerance 3.75 Da for DIA and 2.1 Da for DDA, fragment ion mass tolerance 1 Da (default), two allowed missed cleavages, and one required enzyme specific terminus. Data was searched against the Human International Protein Index (IPI) database version 3.49 which contains a total of 74017 proteins. Identifications were filtered at a ≥ 0.95 protein and peptide probability using PeptideProhpet,[26] and proteins with only one peptide identified were excluded, allowing for a false positive rate of ≤ 0.7%. Identifications resulting from protein isoforms were considered redundant unless a unique peptide was identified to a given secondary isoform. Relative protein levels were determined using spectral counting as previously described using the same ≥ 0.9 probability.[27] Briefly, spectral counting uses the number of peptide tandem mass spectra identified for each peptide from a given protein to estimate abundance of that protein relative to all others present in the sample. To adapt the output of spectral counting for both rational gas phase fractionation and PAcIFIC datasets, the spectral counts for each protein were summed for all data over the complete *m/z* range analyzed per sample for each replicate. The final protein spectral count was the accumulation of all spectral counts from all PAcIFIC, or genome based GPF, HPLC MS runs for one analytical cycle. A t-test was used to determine if the difference between relative levels were statistically significant with a cutoff of greater than 0.05, and of this statistically significant set, identification in both analytical duplicates in either control or tamoxifen-treated was required.

Results

Global Protein Identifications

Total cell lysates from MCF-7 cells +/− tamoxifen treatment were prepared for proteomic profiling as described in the experimental section. Data were acquired in duplicate using two direct, but different, shotgun proteomic discovery methods that eschew sample pre-fractionation: **1)** DDA genome based GPF analysis with the following *m/z* regions 350–520, 515–690, 685–970, and 965–2000 and **2)** DIA PAcIFIC analysis where each *m/z* channel from 400 to 1400 *m/z* (e.g. 400, 401.5, 403, 404.5 …) was examined by CID. For both methods, database search identified proteins were filtered at a minimum probability of 0.95 with two or more peptide tandem mass spectral matches required for positive protein identification. From these experiments the PAcIFIC method identified 1791 proteins in the control sample and 1631 proteins from tamoxifen-treated cells. From the list of proteins identified in the control cells 501 were unique whereas 341 were unique to the tamoxifen treated cell proteome, for a total of 2132 proteins identified from the two conditions. While the analytical efficiency (proteins identified/injection) of genome-based GPF analysis was four-fold greater than that of PAcIFIC, only half as many proteins were identified; i.e. 841 proteins from the control and 785 proteins from tamoxifen-treated cells. Fewer unique proteins were also identified by GPF in the control versus tamoxifen treated cells; i.e. 276 versus 220 respectively. Notably, using the DIA PAcIFIC strategy, we identified more than twice the number of peptides per protein and more than twice as many proteins compared to the genome-based DDA GPF method. While the PAcIFIC method is only 36% as efficient as genome-based GPF in terms of proteins identified per HPLC-MS/MS analysis (i.e. 66 versus 24), the increase in individual protein sequence coverage to an average of three, rather than two peptides/protein, added confidence to quantification by spectral counting discussed below. Additionally, PAcIFIC analysis identified approximately three times more

transcription factors than the previous studies (Umar *et al*) and twice as many as the DDA GPF results described above (Table 1).

Quantitative Analysis

Spectral counting was used on all HPLC-MS/MS data sets to identify differences in relative levels of protein expression between +/− tamoxifen-treated MCF-7 cells (Table 2). Quantitative results from genome-based GPF identified 19 proteins with a statistically significant change in expression in the tamoxifen-treated cells, three of which were also corroborated by PAcIFIC analysis (Supplemental Table 1). The PAcIFIC method, however, identified 83 proteins as differentially expressed at a statistically significant level in the tamoxifen-treated MCF-7 cells with 23 and 60 proteins at higher and lower levels, respectively, compared to control (Table 2). The number of spectral counts per protein was equal to, or significantly higher, for DIA PAcIFIC analysis than the DDA genome based GPF results (Figure 1). The higher spectral counts derived from PAcIFIC then provided a more rigorous statistical analysis for the identification of statistically significant changes in protein abundance.

Target Validation

One of the most up-regulated proteins after tamoxifen treatment identified in this study was AGR2 (Table 2), a recently discovered protein implicated in cell migration, limb regeneration, p53 down-regulation, poor survival rates, and metastasis.[20,28–32] Having observed that a relatively novel pro-oncogenic protein was induced by tamoxifen treatment, we focused biochemical validation on investigating whether AGR2 might function as a tamoxifen agonistic or antagonistic factor. The treatment of cells with tamoxifen showed, as expected, that AGR2 is induced over time using immunochemical methods (Figure 2A), which correlated with elevations in AGR2 mRNA levels using RT-PCR (Figure 2B). However, despite elevated AGR2 mRNA in a microarray study (data not shown), a parallel transcriptomic screen did not identify AGR2 as a significant outlier that would warrant evaluation (Supplemental Table 2).

Common assays used to evaluate the function of a cancer gene involve determining whether the gene can affect cell survival or circumvent drug-induced cell growth inhibition. Using cells stably transfected with an AGR2 gene, or vector control,[20] we demonstrated that the AGR2 gene can promote resistance to cell death induced by the common DNA damaging agent doxorubicin (Figure 3). Further, although these two cell panels (vector control and AGR2) do not show obvious differences in the growth rate in xenograft animal models, the cells expressing the AGR2 gene confer resistance to cisplatin (Figure 4). Thus, by demonstrating that AGR2 can mediate drug resistance using *in vitro* and *in vivo* xenograft model systems, we further validated this PAcIFIC generated hypothesis.

Discussion

Tamoxifen has provided substantial improvements in disease management, however it has also been associated with agonistic effects involving either intrinsic, or acquired, resistance. Previous studies comparing tamoxifen-induced proteomes have primarily used 2-DE methods with subsequent HPLC-MS/MS, or matrix assisted laser desorption ionization mass spectrometry (MALDI), for identification of proteins present in each stained spot.[33–34] As expected, 2-DE methods tend to identify fewer proteins than shotgun proteomic methods, in part, because of loading capacity issues with the first dimension. Recently, the first paper using shotgun proteomics to characterize tissues from tamoxifen-sensitive and tamoxifen-resistant tumors using an accurate mass and time (AMT) tag approach was published.[35] The AMT shotgun proteomic method consists of two phases. In the first phase, an extensive

library of proteins identified by DDA-based HPLC-MS/MS of peptides is recorded in a library according to mass and elution time using any available material. In the second phase, conducted on samples of limited abundance using only precursor ion scans, the AMT peptide library of previously identified proteins is used to identify proteins by comparison to the library. Thus, the method can be quite sensitive and this particular AMT study identified 2556 proteins from 51 tumor samples whereas the prior 2-DE studies identify only between 12–20 proteins.

In an effort to analyze proteomic samples more efficiently, our laboratory has focused recently on developing shotgun proteomic methods that eschew fractionation prior to HPLC-MS analysis. Specifically, one of these methods, PAcIFIC, has been shown to be capable of detecting proteins an order of magnitude lower in abundance than standard methods that fractionate peptides prior to DDA HPLC-MS/MS analysis (Panchaud *et al*., 2009) and to detect proteins across the full dynamic range of yeast (Panchaud *et al*., 2011).[36] Here PAcIFIC was employed to identify tamoxifen-induced proteins for the evaluation of agonistic or antagonistic effects. In this study, we used our prior optimized DDA genome-based GPF method (Scherl *et al*., 2008) and the newer DIA PAcIFIC method to identify proteins in total cell lysates from MCF-7 cells +/− tamoxifen treatment. While the analytical efficiency of proteins identified per HPLC-MS/MS analysis is four-fold greater for the genome based GPF versus the PAcIFIC method, the number of proteins identified reaches a plateau beyond which additional analyses do not result in further proteins being identified.[9] As we have shown, this is most likely due to inefficient DDA sampling and, in fact, previous results demonstrated that the GPF method worsened as the *m/z* range sampled was reduced to ~ 50 *m/z*, suggesting a fundamental problem with DDA sampling from too narrow an *m/z* range that prevented further refinement of this approach.[9,37] This led to development of the PAcIFIC method that systematically queries peptides present in all *m/z* channels without regard for the presence/absence of a precursor ion. While this feature reduces analytical efficiency, it extends dynamic range to a group of peptides not detectable by DDA methods because they are obscured by baseline chemical noise or high intensity ion signals.

The lower efficiency of PAcIFIC analysis indicates that sampling has reached the bottom of available mass spectrometric dynamic range; e.g. depending on complexity between 15–30% of peptides identified can come from peptides with no detectable precursor ion which we previously defined as orphan peptides (Panchaud *et al*., 2009). The importance of more complete proteomic profiling provided by PAcIFIC is exemplified by the case of AGR2 which was identified in both the DDA GPF and DIA PAcIFIC datasets, but only PAcIFIC detected a significant change in expression. This detection of AGR2 as having a significant change in expression post tamoxifen treatment by PAcIFIC was due to the higher spectral count values in the PAcIFIC data than were detected in the GPF data. This highlights an additional benefit of the PAcIFIC method which detects more peptides per protein (Table 1) than the DDA GPF method and thus allows for more accurate quantification by spectral counting (Table 3) and extension of the detectable quantitative dynamic range. Finally, we note that when sample is limited (e.g. < 24 HPLC-MS/MS runs are possible) that use of the genome based GPF method is recommended over PAcIFIC analysis because of the higher efficiency.

In our DIA PAcIFIC data, but not the DDA GPF data, one of the most significantly induced proteins, AGR2, stood out as a recently identified pro-oncogenic protein. Given the growing interest in AGR2 in cancer research, we included a novel biological validation to affirm the importance of AGR2 and/or its pathway partners as a potential anti-cancer drug target. This investigation into the importance of the up-regulation of AGR2 on cell viability and tamoxifen agonistic effects demonstrated that AGR2 is directly linked to cell viability and that tamoxifen induction of AGR2 protein occurs at the transcriptional level (Figures 2a and

2b). For example, it has been demonstrated previously that AGR2 can induce metastasis in animal models and mediate limb regeneration.[28–31] AGR2 can also predict poor prognosis when over-expressed in prostate and breast cancers.[32,38] Furthermore, the AGR2 gene can function as a drug resistance factor in model cancer systems (Figure 3 and 4), suggesting that it may provide agonistic effects from tamoxifen if it were induced *in vivo* in response to the drug.

Of the other proteins identified at elevated levels after tamoxifen treatment in the PAcIFIC dataset (Table 2), several besides AGR2 are associated with cancer cell growth. One of the most up-regulated proteins, Aminopeptidase B, catalyzes the hydrolysis of amino acids, and inhibition of it would disrupt protein degradation, interfering with cell turnover. Research studies on the inhibition of aminopeptidases as cancer therapeutics have already begun in myeloma cell lines.[39] Clusterin precursor was also identified as up-regulated due to tamoxifen treatment, and recent studies demonstrate that inhibition of clusterin, a pro-apoptotic protein, increased the sensitivities to tamoxifen treatment and is at higher levels in cell lines resistant to anti-estrogen treatments.[40–41] Septins have also been associated with cancer and Septin 2, identified as up-regulated post tamoxifen treatment in this study, is required for cytokinesis and binds to actin. For example, Septin 9 is much more abundant in breast cancer cells than in normal cells, and has been shown to enhance cell growth and motility.[42] Finally, we observed an overall up-regulation of proteins associated with the cytoskeleton and cell motility in our data, perhaps suggesting a link between tamoxifen treatment, AGR2, Septins and cancer.

Despite the differences in sample source and analytical methods used to acquire data, there was a 60% overlap between proteins identified in the +/− tamoxifen-treated MCF-7 cells in the present study and the clinical samples of Umar *et al*. We feel this, as well as the fact that PAcIFIC profiling identified almost three times as many transcription factors (Table 1) compared to the number identified in the DDA-based Umar study[35], provide further evidence in favor of the DIA-based proteome profiling. Additionally, we note that AGR2, a protein we detected and confirmed a pro-survival role in our study, was not identified in the Umar AMT study. Regardless, we believe that the two studies provide complementary information on tamoxifen biochemistry due to **i)** the differences between heterogeneous clinical samples and homogeneous cell lines as well as the differences between the treatment regimens of patients and selectively treated cell lines, and **ii)** differences in the analytical methods used, which identified roughly the same number of proteins. Importantly, because the DIA PAcIFIC method screens for peptides in each *m/z* channel, regardless of precursor ion signal, effective dynamic range is extended. Thus, PAcIFIC could be used to extend the dynamic range of AMT databases because the classic DDA approach used to build an AMT library selects only peptide ions that are above the chemical noise threshold limiting the dynamic range of detection. Illustrating this point, we note that the Umar *et al*. AMT study on clinical samples, did not identify AGR2 as an important regulator of tamoxifen treatment, suggesting that **i)** the DDA methods used to create the AMT database do not probe a proteome as deeply as the DIA PAcIFIC method and/or **ii)** there is a role for proteomic analysis of more homogeneous cancer cell lines alongside heterogeneous clinical studies. Regardless of these deficiencies, the AMT method excels at analysis of clinical samples of limited availability because proteins are identified from a single HPLC-MS experiment through comparison to the predefined AMT peptide library. The work in this manuscript required 45 μg of protein, 1 μg for each of the 45 LCMS acquisitions, per sample for analysis; however, since then the PAcIFIC method has been further optimized. Specifically, the *m/z* range covered and the number of channels in each acquisition were increased due to the faster scan speeds of the LTQ Velos, effectively reducing the number of LCMS acquisitions and amount of protein required for a complete PAcIFIC analysis.[36] Currently, a full PAcIFIC analysis requires only ~ 20 μg of protein, which is more similar to the number

needed to set up an AMT database than to profile clinical samples, but this number will fall as new instrumentation is developed and ion trap scan speeds increase; e.g. the LTQ-Velos scans roughly twice as fast as the prior model by incorporating an ion funnel prior to the ion trap.[43–46] Thus, we suggest that a future clinical approach to proteome profiling would combine the advantages of limited sample requirements to read identifications from an AMT database with the greater detectable dynamic range provided by PAcIFIC. Finally, we note that similar to building an AMT database, SRM databases which are used to provide quantitative data by following specific peptide fragment ion transitions are available directly from any given PAcIFIC data set for most, if not all, peptides present in a sample.[47]

In summary, cancer is proving to be a tissue-specific disease and novel approaches are required to improve drug responses and disease management in patients. While most of the current highly sensitive proteomic approaches can be used to identify differential protein expression in drug-resistant and drug-sensitive clinical cancers to identify putative drug-resistant or prognostic factors, they have a limited detectable dynamic range by use of the DDA process that fails to select most of the low intensity ion signals. As we have shown here, use of a simple DIA proteomic profiling approach, PAcIFIC, extended dynamic range of detection allowing an important pro-survival protein, AGR2, to be detected as up-regulated in MCF-7 cells post tamoxifen treatment, a fact we then confirmed in cell based assays. AGR2 has also been implicated in other types of cancer, including pancreatic and prostate, where AGR2 is thought to play a role in predicting poor prognosis and metastatic growth. Furthermore, the lack of identification of AGR2, known to be important in tamoxifen response, at statistically significant levels using transcriptomic and standard DDA shotgun proteomic methods highlights the need for continued development of more thorough proteome profiling methods. We note that, while the ion trap instrumentation used in this study required two days of continuous MS use for a single PAcIFIC experiment, that the use of ion funnels to decrease ion trap fill times (e.g. the LTQ Velos, ThermoFisher) now allows the same experiment to be conducted in only 1.5 days. We expect that this trend toward even faster scanning mass spectrometers will continue and from this trend more thorough proteome screens that eventually rival transcriptomic screens for speed and sensitivity will emerge. Thus, we feel that the simplicity and directness of PAcIFIC holds promise for more routine screening of primary clinical material that would complement whole genome transcriptome and whole genome cancer sequencing.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Abbreviations

AGC	Automatic Gain Control
AGR2	Anterior gradient homolog 2 protein
AMT	Accurate Mass and Time tag
DDA	Data-Dependent Acquisition
DIA	Data-Independent Analysis
GPF	Gas Phase Fractionation
IPI	International Protein Index
LTQ	Linear Ion Trap
PAcIFIC	Precursor Acquisition Independent From Ion Count

TMT	Tandem Mass Tags
SILAC	Stable Isotope Labeling with Amino acids in cell Culture
2-DE	Two Dimensional Electrophoresis

Acknowledgments

The authors wish to thank the U.S.A. NIH Pharmaceutical Sciences Training Grant (T32 GM07750), U.S.A. NIH NCRR (1S10RR017262-01), the University of Washington's Proteomics Resource (UWPR95794), Cancer Research UK (C483/A6354), and BBSRC (UK) (BB/C511599/1) for funding and support. We also wish to thank Jason Shaw and Dr. Brook L. Nunn for reviewing this manuscript, and Ying (Sonia) Ting for technical assistance.

References

1. Mardis ER, Ding L, Dooling DJ, Larson DE, McLellan MD, Chen K, Koboldt DC, Fulton RS, Delehaunty KD, McGrath SD, Fulton LA, Locke DP, Magrini VJ, Abbott RM, Vickery TL, Reed JS, Robinson JS, Wylie T, Smith SM, Carmichael L, Eldred JM, Harris CC, Walker J, Peck JB, Du F, Dukes AF, Sanderson GE, Brummett AM, Clark E, McMichael JF, Meyer RJ, Schindler JK, Pohl CS, Wallis JW, Shi X, Lin L, Schmidt H, Tang Y, Haipek C, Wiechert ME, Ivy JV, Kalicki J, Elliott G, Ries RE, Payton JE, Westervelt P, Tomasson MH, Watson MA, Baty J, Heath S, Shannon WD, Nagarajan R, Link DC, Walter MJ, Graubert TA, DiPersio JF, Wilson RK, Ley TJ. Recurring mutations found by sequencing an acute myeloid leukemia genome. N Engl J Med. 2009; 361(11): 1058–66. [PubMed: 19657110]
2. Mardis ER, Wilson RK. Cancer genome sequencing: a review. Hum Mol Genet. 2009; 18(R2):R163–R168. [PubMed: 19808792]
3. Pleasance ED, Cheetham RK, Stephens PJ, McBride DJ, Humphray SJ, Greenman CD, Varela I, Lin ML, Ordonez GR, Bignell GR, Ye K, Alipaz J, Bauer MJ, Beare D, Butler A, Carter RJ, Chen L, Cox AJ, Edkins S, Kokko-Gonzales PI, Gormley NA, Grocock RJ, Haudenschild CD, Hims MM, James T, Jia M, Kingsbury Z, Leroy C, Marshall J, Menzies A, Mudie LJ, Ning Z, Royce T, Schulz-Trieglaff OB, Spiridou A, Stebbings LA, Szajkowski L, Teague J, Williamson D, Chin L, Ross MT, Campbell PJ, Bentley DR, Futreal PA, Stratton MR. A comprehensive catalogue of somatic mutations from a human cancer genome. Nature. 463(7278):191–196. [PubMed: 20016485]
4. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. Nature. 2009; 458(7239):719–724. [PubMed: 19360079]
5. Nie L, Wu G, Zhang W. Correlation between mRNA and protein abundance in Desulfovibrio vulgaris: a multiple regression to identify sources of variations. Biochem Biophys Res Commun. 2006; 339(2):603–610. [PubMed: 16310166]
6. Chen G, Gharib TG, Huang CC, Taylor JM, Misek DE, Kardia SL, Giordano TJ, Iannettoni MD, Orringer MB, Hanash SM, Beer DG. Discordant protein and mRNA expression in lung adenocarcinomas. Mol Cell Proteomics. 2002; 1(4):304–313. [PubMed: 12096112]
7. Washburn MP, Wolters D, Yates JR 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol. 2001; 19(3):242–247. [PubMed: 11231557]
8. Neville DC, Rozanas CR, Price EM, Gruis DB, Verkman AS, Townsend RR. Evidence for phosphorylation of serine 753 in CFTR using a novel metal-ion affinity resin and matrix-assisted laser desorption mass spectrometry. Protein Sci. 1997; 6(11):2436–2445. [PubMed: 9385646]
9. Scherl A, Shaffer SA, Taylor GK, Kulasekara HD, Miller SI, Goodlett DR. Genome-specific gas-phase fractionation strategy for improved shotgun proteomic profiling of proteotypic peptides. Anal Chem. 2008; 80(4):1182–1191. [PubMed: 18211032]
10. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics. 2002; 1(5):376–386. [PubMed: 12118079]

11. Choe LH, Aggarwal K, Franck Z, Lee KH. A comparison of the consistency of proteome quantitation using two-dimensional electrophoresis and shotgun isobaric tagging in Escherichia coli cells. Electrophoresis. 2005; 26(12):2437–2449. [PubMed: 15924362]
12. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet-Jones M, He F, Jacobson A, Pappin DJ. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics. 2004; 3(12):1154–1169. [PubMed: 15385600]
13. Thompson A, Schafer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, Neumann T, Johnstone R, Mohammed AK, Hamon C. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal Chem. 2003; 75(8):1895–1904. [PubMed: 12713048]
14. Panchaud A, Scherl A, Shaffer SA, von Haller PD, Kulasekara HD, Miller SI, Goodlett DR. Precursor Acquisition Independent From Ion Count: How to Dive Deeper into the Proteomics Ocean. Anal Chem. 2009
15. Veronesi U, Boyle P, Goldhirsch A, Orecchia R, Viale G. Breast cancer. Lancet. 2005; 365(9472): 1727–41. [PubMed: 15894099]
16. Hinestrosa MC, Dickersin K, Klein P, Mayer M, Noss K, Slamon D, Sledge G, Visco FM. Shaping the future of biomarker research in breast cancer to ensure clinical relevance. Nat Rev Cancer. 2007; 7(4):309–315. [PubMed: 17384585]
17. Clarke R, Liu MC, Bouker KB, Gu Z, Lee RY, Zhu Y, Skaar TC, Gomez B, O'Brien K, Wang Y, Hilakivi-Clarke LA. Antiestrogen resistance in breast cancer and the role of estrogen receptor signaling. Oncogene. 2003; 22(47):7316–7339. [PubMed: 14576841]
18. Massarweh S, Schiff R. Resistance to endocrine therapy in breast cancer: exploiting estrogen receptor/growth factor signaling crosstalk. Endocr Relat Cancer. 2006; (13 Suppl 1):S15–S24. [PubMed: 17259554]
19. Cory AH, Owen TC, Barltrop JA, Cory JG. Use of an aqueous soluble tetrazolium/formazan assay for cell growth assays in culture. Cancer Commun. 1991; 3(7):207–212. [PubMed: 1867954]
20. Pohler E, Craig AL, Cotton J, Lawrie L, Dillon JF, Ross P, Kernohan N, Hupp TR. The Barrett's antigen anterior gradient-2 silences the p53 transcriptional response to DNA damage. Mol Cell Proteomics. 2004; 3(6):534–547. [PubMed: 14967811]
21. Sewell JM, Macleod KG, Ritchie A, Smyth JF, Langdon SP. Targeting the EGF receptor in ovarian cancer with the tyrosine kinase inhibitor ZD 1839 ("Iressa"). Br J Cancer. 2002; 86(3):456–462. [PubMed: 11875715]
22. Macleod K, Mullen P, Sewell J, Rabiasz G, Lawrie S, Miller E, Smyth JF, Langdon SP. Altered ErbB receptor signaling and gene expression in cisplatin-resistant ovarian cancer. Cancer Res. 2005; 65(15):6789–6800. [PubMed: 16061661]
23. Workman P, Aboagye EO, Balkwill F, Balmain A, Bruder G, Chaplin DJ, Double JA, Everitt J, Farningham DA, Glennie MJ, Kelland LR, Robinson V, Stratford IJ, Tozer GM, Watson S, Wedge SR, Eccles SA. Guidelines for the welfare and use of animals in cancer research. Br J Cancer. 2010; 102(11):1555–1577. [PubMed: 20502460]
24. Yi EC, Lee H, Aebersold R, Goodlett DR. A microcapillary trap cartridge-microcapillary high-performance liquid chromatography electrospray ionization emitter device capable of peptide tandem mass spectrometry at the attomole level on an ion trap mass spectrometer with automated routine operation. Rapid Commun Mass Spectrom. 2003; 17(18):2093–2098. [PubMed: 12955739]
25. Eng JK, Mccormack AL, Yates JR. An Approach to Correlate Tandem Mass-Spectral Data of Peptides with Amino-Acid-Sequences in a Protein Database. Journal of the American Society for Mass Spectrometry. 1994; 5(11):976–989.
26. Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem. 2003; 75(17):4646–4658. [PubMed: 14632076]
27. Ryu S, Gallis B, Goo YA, Shaffer SA, Radulovic D, Goodlett DR. Comparison of a label-free quantitative proteomic method based on Peptide ion current area to the isotope coded affinity tag method. Cancer Inform. 2008; 6:243–255. [PubMed: 19259412]

28. Ramachandran V, Arumugam T, Wang H, Logsdon CD. Anterior gradient 2 is expressed and secreted during the development of pancreatic cancer and promotes cancer cell survival. Cancer Res. 2008; 68(19):7811–7818. [PubMed: 18829536]
29. Innes HE, Liu D, Barraclough R, Davies MP, O'Neill PA, Platt-Higgins A, de Silva Rudland S, Sibson DR, Rudland PS. Significance of the metastasis-inducing protein AGR2 for outcome in hormonally treated breast cancer patients. Br J Cancer. 2006; 94(7):1057–1065. [PubMed: 16598187]
30. Kumar A, Godwin JW, Gates PB, Garza-Garcia AA, Brockes JP. Molecular basis for the nerve dependence of limb regeneration in an adult vertebrate. Science. 2007; 318(5851):772–7. [PubMed: 17975060]
31. Wang Z, Hao Y, Lowe AW. The adenocarcinoma-associated antigen, AGR2, promotes tumor growth, cell migration, and cellular transformation. Cancer Res. 2008; 68(2):492–497. [PubMed: 18199544]
32. Barraclough DL, Platt-Higgins A, de Silva Rudland S, Barraclough R, Winstanley J, West CR, Rudland PS. The Metastasis-Associated Anterior Gradient 2 Protein Is Correlated with Poor Survival of Breast Cancer Patients. Am J Pathol. 2009
33. Neubauer H, Clare SE, Kurek R, Fehm T, Wallwiener D, Sotlar K, Nordheim A, Wozny W, Schwall GP, Poznanovic S, Sastri C, Hunzinger C, Stegmann W, Schrattenholz A, Cahill MA. Breast cancer proteomics by laser capture microdissection, sample pooling, 54-cm IPG IEF, and differential iodine radioisotope detection. Electrophoresis. 2006; 27(9):1840–1852. [PubMed: 16645950]
34. Wang Y, He QY, Chen H, Chiu JF. Synergistic effects of retinoic acid and tamoxifen on human breast cancer cells: proteomic characterization. Exp Cell Res. 2007; 313(2):357–368. [PubMed: 17098229]
35. Umar A, Kang H, Timmermans AM, Look MP, Meijer-van Gelder ME, den Bakker MA, Jaitly N, Martens JW, Luider TM, Foekens JA, Pasa-Tolic L. Identification of a putative protein profile associated with tamoxifen therapy resistance in breast cancer. Mol Cell Proteomics. 2009; 8(6): 1278–1294. [PubMed: 19329653]
36. Panchaud A, Jung S, Shaffer SA, Aitchison JD, Goodlett DR. Faster, quantitative, and accurate precursor acquisition independent from ion count. Anal Chem. 2011; 83(6):2250–2257. [PubMed: 21341720]
37. Yi EC, Marelli M, Lee H, Purvine SO, Aebersold R, Aitchison JD, Goodlett DR. Approaching complete peroxisome characterization by gas-phase fractionation. Electrophoresis. 2002; 23(18): 3205–3216. [PubMed: 12298092]
38. Zhang Y, Forootan SS, Liu D, Barraclough R, Foster CS, Rudland PS, Ke Y. Increased expression of anterior gradient-2 is significantly associated with poor survival of prostate cancer patients. Prostate Cancer Prostatic Dis. 2007; 10(3):293–300. [PubMed: 17457305]
39. Moore HE, Davenport EL, Smith EM, Muralikrishnan S, Dunlop AS, Walker BA, Krige D, Drummond AH, Hooftman L, Morgan GJ, Davies FE. Aminopeptidase inhibition as a targeted treatment strategy in myeloma. Mol Cancer Ther. 2009; 8(4):762–770. [PubMed: 19372548]
40. Redondo M, Tellez T, Roldan MJ, Serrano A, Garcia-Aranda M, Gleave ME, Hortas ML, Morell M. Anticlusterin treatment of breast cancer cells increases the sensitivities of chemotherapy and tamoxifen and counteracts the inhibitory action of dexamethasone on chemotherapy-induced cytotoxicity. Breast Cancer Res. 2007; 9(6):R86. [PubMed: 18078515]
41. Toffanin S, Daidone MG, Miodini P, De Cecco L, Gandellini P, Cappelletti V. Clusterin: a potential target for improving response to antiestrogens. Int J Oncol. 2008; 33(4):791–798. [PubMed: 18813793]
42. Gonzalez ME, Peterson EA, Privette LM, Loffreda-Wren JL, Kalikin LM, Petty EM. High SEPT9_v1 expression in human breast cancer cells is associated with oncogenic phenotypes. Cancer Res. 2007; 67(18):8554–8564. [PubMed: 17875694]
43. Shaffer SA, Prior DC, Anderson GA, Udseth HR, Smith RD. An ion funnel interface for improved ion focusing and sensitivity using electrospray ionization mass spectrometry. Anal Chem. 1998; 70(19):4111–4119. [PubMed: 9784749]

44. Shaffer SA, Tolmachev A, Prior DC, Anderson GA, Udseth HR, Smith RD. Characterization of an improved electrodynamic ion funnel interface for electrospray ionization mass spectrometry. Anal Chem. 1999; 71(15):2957–2964. [PubMed: 10450147]

45. Olsen JV, Schwartz JC, Griep-Raming J, Nielsen ML, Damoc E, Denisov E, Lange O, Remes P, Taylor D, Splendore M, Wouters ER, Senko M, Makarov A, Mann M, Horning S. A dual pressure linear ion trap orbitrap instrument with very high sequencing speed. Mol Cell Proteomics. 2009; 8(12):2759–2769. [PubMed: 19828875]

46. Second TP, Blethrow JD, Schwartz JC, Merrihew GE, MacCoss MJ, Swaney DL, Russell JD, Coon JJ, Zabrouskov V. Dual-pressure linear ion trap mass spectrometer improving the analysis of complex protein mixtures. Anal Chem. 2009; 81(18):7757–7765. [PubMed: 19689114]

47. Picotti P, Rinner O, Stallmach R, Dautel F, Farrah T, Domon B, Wenschuh H, Aebersold R. High-throughput generation of selected reaction-monitoring assays for proteins and proteomes. Nat Methods. 2010; 7(1):43–46. [PubMed: 19966807]

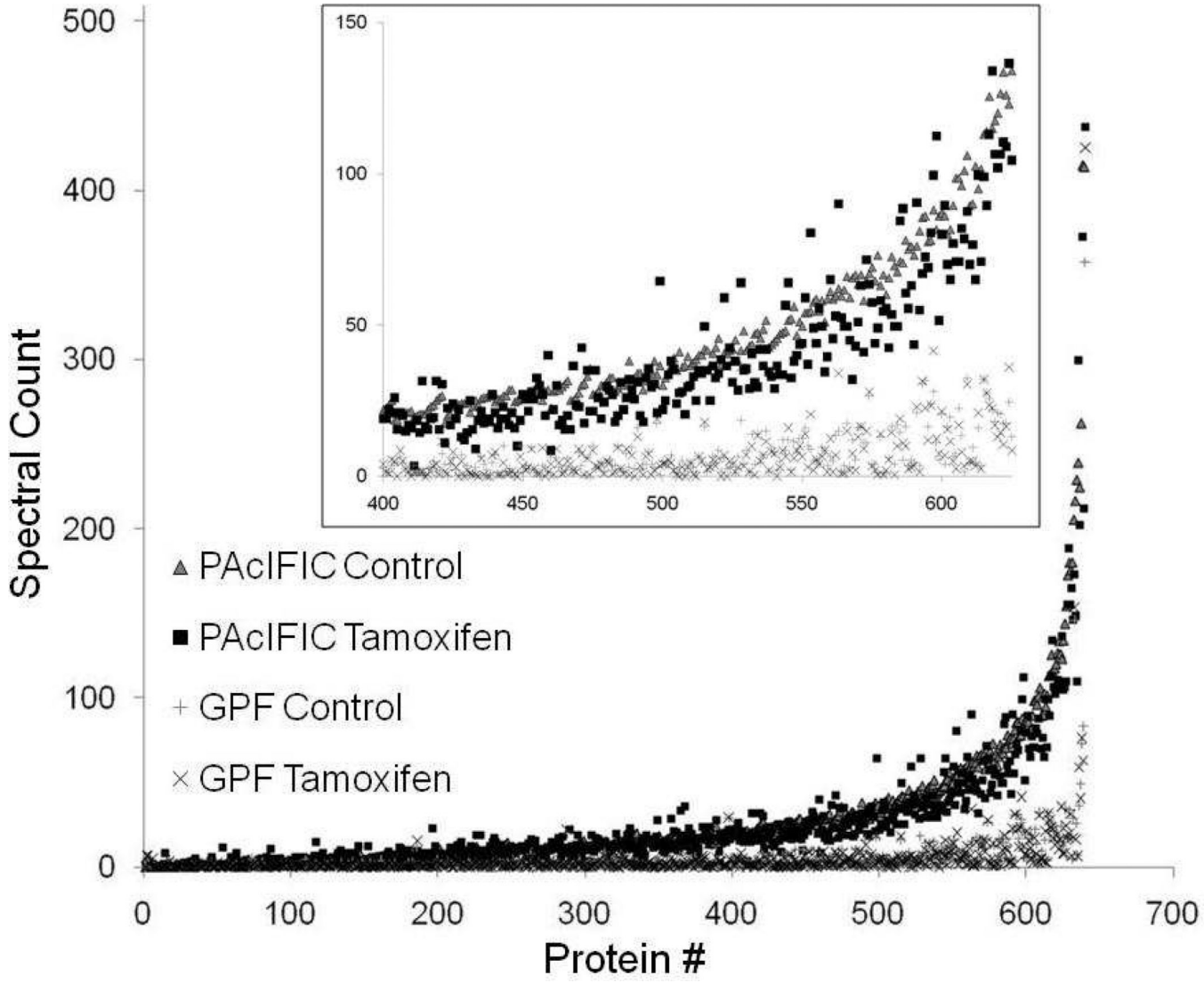

Figure 1.
Spectral count per protein with DDA GPF vs PAcIFIC analysis for all mutual protein identifications.

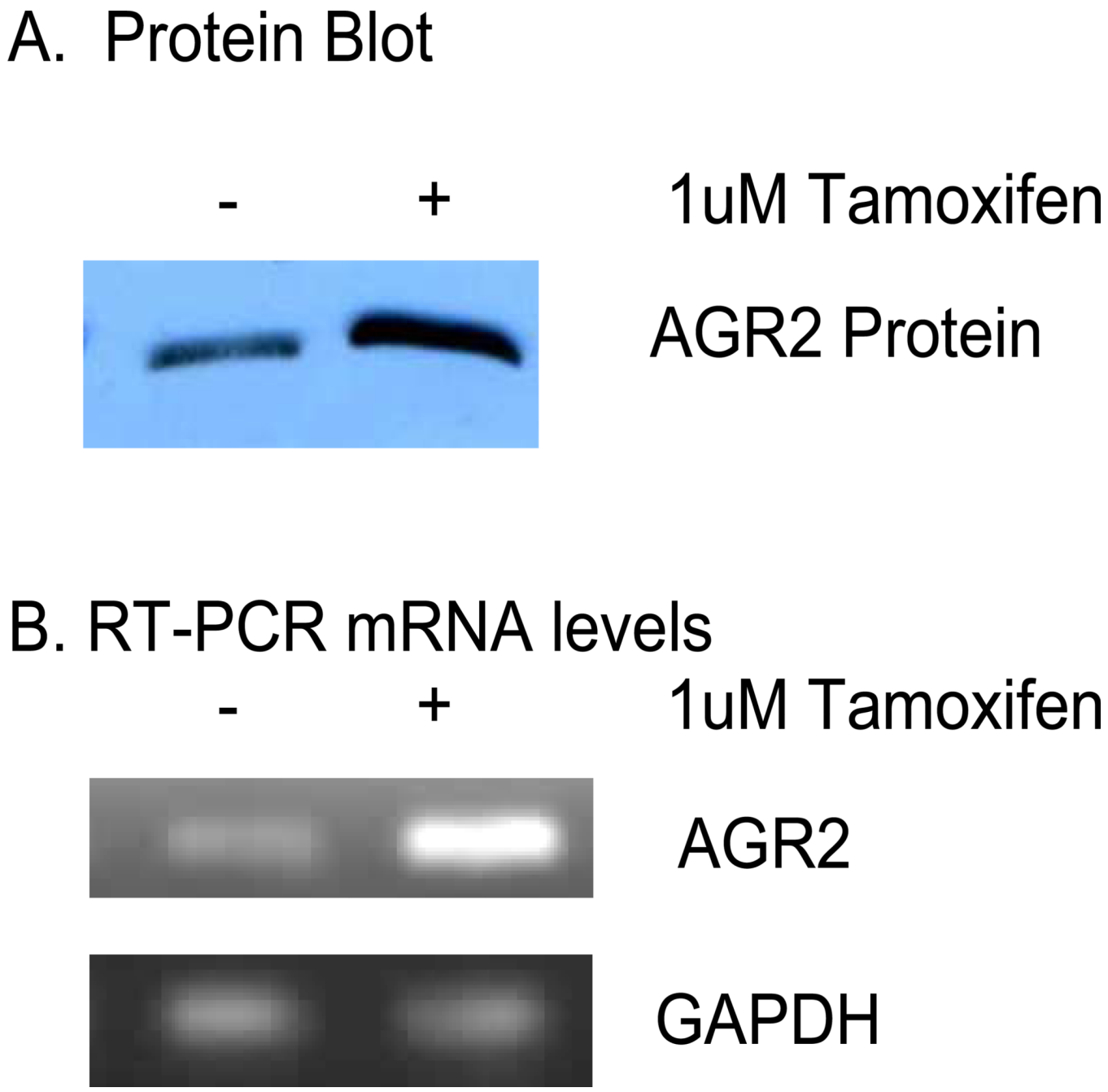

Figure 2.
Induction of AGR2 by tamoxifen. A) Protein lysate immunoblotted for AGR2 protein in response to tamoxifen treatment B) AGR2 mRNA levels in response to tamoxifen treatment (GAPDH control).

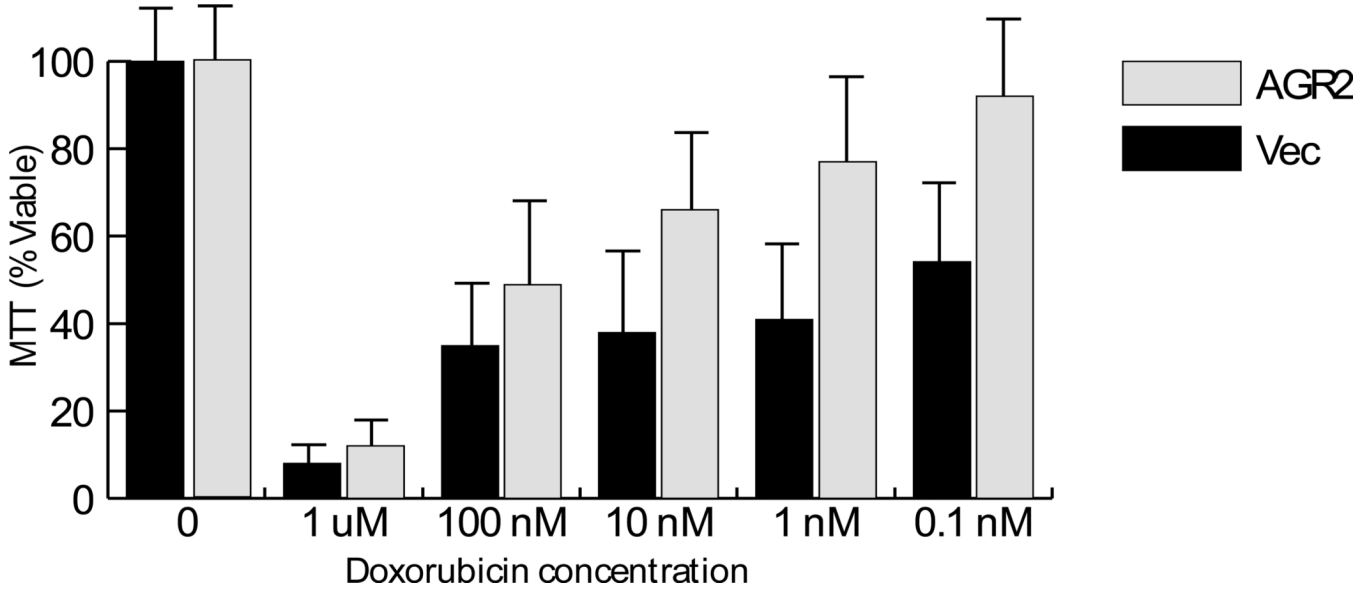

Figure 3.
AGR2 mediates in vitro resistance to doxorubicin. H1299 cells expressing AGR2 or control cells were incubated with increasing concentrations of doxorubicin and processed for cell viability using the MTT assay.

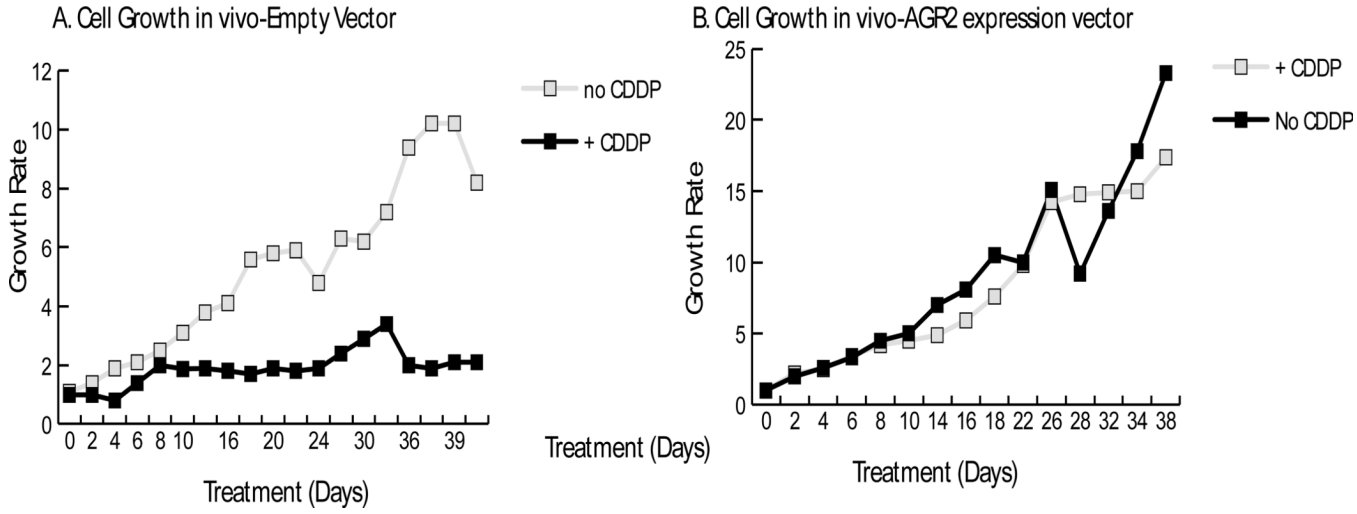

Figure 4.

AGR2 Mediates Resistance of Xenograft cells to cisplatin administration. Isogenic cells with vector control (A), or stably expressed AGR2 (B), were grown in xenograft systems and cisplatin (CDDP) was administered over the indicated time course.

Table 1

Number of peptides identified and sequence coverage per transcription factor identified in PAcIFIC and genome based GPF analysis.

IPI	PAcIFIC % Sequence Coverage	GPF % Sequence Coverage	PAcIFIC Unique Peptides	GPF Unique Peptides	PAcIFIC Total Peptides	GPF Total Peptides	Gene Symbol*	Description
IPI00005184.1	9.4	2.2	5	1	4	2	TRIM24	TRIM24 Isoform Long of Transcription intermediary factor 1-alpha
IPI00006079.1	10.7	-	6	-	3	-	BCLAF1	BCLAF1 Isoform 1 of Bcl-2-associated transcription factor 1
IPI00012901.1	14.2	10.8	4	3	4	3	GATA3	GATA3 Isoform 1 of Trans-acting T-cell-specific transcription factor GATA-3
IPI00014533.1	11.8	5.2	6	3	4	3	UBTF	UBTF Isoform UBF1 of Nucleolar transcription factor 1
IPI00015806.3	19.1	-	4	-	2	-	GTF3C3	GTF3C3 Isoform 1 of General transcription factor 3C polypeptide 3
IPI00016725.1	15.5	-	8	-	6	-	GTF3C4	GTF3C4 General transcription factor 3C polypeptide 4
IPI00020928.1	28.5	29.4	6	7	4	9	TFAM	TFAM Transcription factor A
IPI00022348.1	8.9	-	3	-	2	-	PML	PML Isoform PML-1 of Probable transcription factor PML
IPI00030781.1	18.5	2.2	10	1	10	4	STAT1	STAT1 Isoform Alpha of Signal transducer and activator of transcription 1-alpha/ beta
IPI00054042.1	32.9	17.9	24	11	18	19	GTF2I	GTF2I Isoform 1 of General transcription factor II-I
IPI00170594.5	6.6	-	9	-	8	-	AHCTF1	AHCTF1 Isoform 1 of AT-hook-containing transcription factor 1
IPI00220109.6	6.3	-	9	-	7	-	ATRX	ATRX transcriptional regulator ATRX isoform 2
IPI00221035.4	29	11.7	2	1	2	3	BTF3	BTF3 Isoform 1 of Transcription factor BTF3
IPI00221222.7	26	15.7	4	2	4	4	SUB1	SUB1 Activated RNA polymerase II transcriptional coactivator p15
IPI00247871.2	6.1	3.2	5	2	3	2	TCERG1	TCERG1 Isoform 1 of Transcription elongation regulator 1
IPI00254408.6	4	0.9	7	1	6	5	BPTF	BPTF bromodomain PHD finger transcription factor isoform 1
IPI00296078.5	11.3	-	8	-	7	-	TRPS1	TRPS1 Zinc finger transcription factor Trps1, Gene_Symbol=TRPS1 zinc finger transcription factor TRPS1
IPI00298058.1	7.2	-	5	-	4	-	SUPT5H	SUPT5H Isoform 1 of Transcription elongation factor SPT5
IPI00300341.5	34.4	-	3	-	3	-	TCEB1	TCEB1 Transcription elongation factor B polypeptide 1
IPI00301139.5	4.8	-	3	-	2	-	MED17	MED17 Isoform 1 of Mediator of RNA polymerase II transcription subunit 17
IPI00333215.1	23.9	8.6	4	2	2	2	TCEA1	TCEA1 Isoform 1 of Transcription elongation factor A protein 1
IPI00411531.3	10.4	-	3	-	2	-	GTF3C5	GTF3C5 Isoform 2 of General transcription factor 3C polypeptide 5
IPI00438229.2	60.4	53.7	41	40	34	98	TRIM28	TRIM28 Isoform 1 of Transcription intermediary factor 1-beta
IPI00719549.2	44.2	44.5	5	5	7	11	RBM14	RBM14 Transcriptional coactivator CoAZ
IPI00784161.1	10.3	-	8	-	2	-	SUPT6H	SUPT6H Isoform 1 of Transcription elongation factor SPT6

*Gene Symbol from associated IPI

Table 2

PAcIFIC protein identifications with a statistically different level of expression in tamoxifen treated cells.

IPI	Protein Description	Gene Symbol	Subcellular Localization	Average Spec-Count Control	STDev	Average Spec-Count Tamoxifen	STDev	ttest	Tamoxifen/ Control*
IPI00007427.2	AGR2	AGR2	Secreted	1	1.41	11.5	0.71	0.027	11.50
IPI00642211.3	RNPEP Aminopeptidase B	RNPEP	Secreted	1	0.00	9.5	0.71	0.037	9.50
IPI00291262.3	CLU Clusterin precursor	CLU	Secreted	1.5	0.71	9	1.41	0.045	6.00
IPI00002460.3	-	ANXA7	-	1.5	0.71	5.5	0.71	0.030	3.67
IPI00014177.3	SEPT2 Septin-2	2-Sep	Cytoplasm/Cytoskeleton	3.5	0.71	11	1.41	0.045	3.14
IPI00295252.2	PREX1 Isoform 1 of Phosphatidylinositol 3,4,5-trisphosphate-dependent Rac exchanger 1 protein	PREX1	Cytoplasm	5.5	0.71	17	0.00	0.028	3.09
IPI00002406.2	BCAM Lutheran blood group glycoprotein precursor	BCAM	Membrane	8.5	0.71	22.5	0.71	0.003	2.65
IPI00029623.1	PSMA6 Proteasome subunit alpha type 6	PSMA6	Cytoplasm/Nucleus	2.5	0.71	6.5	0.71	0.030	2.60
IPI00456969.1	DYNC1H1 Dynein heavy chain, cytosolic	DYNC1H1	Cytoplasm	16.5	2.12	35	2.83	0.022	2.12
IPI00298994.6	-	TLN1	Cytoplasm	5.5	0.71	11.5	0.71	0.014	2.09
IPI00010154.3	GDI1 Rab GDP dissociation inhibitor alpha	GDI1	Cytoplasm	5.5	0.71	10.5	0.71	0.019	1.91
IPI00219678.3	EIF2S1 Eukaryotic translation initiation factor 2 subunit 1	EIF2S1	-	4.5	0.71	8.5	0.71	0.030	1.89
IPI00335168.9	MYL6 Isoform Non-muscle of Myosin light polypeptide 6	MYL6; MYL6B	-	12	1.41	22	1.41	0.019	1.83
IPI00028091.3	ACTR3 Actin-like protein 3	ACTR3	Cytoskeleton	8.5	0.71	14.5	0.71	0.014	1.71
IPI00009342.1	IQGAP1 Ras GTPase-activating-like protein IQGAP1	IQGAP1	Cell membrane	17.5	2.12	28	1.41	0.038	1.60
IPI00000816.1	YWHAE 14-3-3 protein epsilon	YWHAE	Cytoplasm	18	1.41	28.5	0.71	0.027	1.58
IPI00218343.4	TUBA1C Tubulin alpha-1C chain	TUBA1C	-	18.5	2.12	28.5	2.12	0.042	1.54
IPI00185374.4	PSMD12 26S proteasome non-ATPase regulatory subunit 12	PSMD12	-	8.5	0.71	12.5	0.71	0.030	1.47
IPI00784156.1	AP2B1 Isoform 1 of AP-2 complex subunit beta-1	AP2B1	Cell membrane	8.5	0.71	12.5	0.71	0.030	1.47
IPI00025512.2	HSPB1 Heat-shock protein beta-1	HSPB1	Cytoplasm/Nucleus	62	1.41	90	2.83	0.018	1.45
IPI00219005.3	FKBP4 FK506-binding protein 4	FKBP4	Cytoplasm/Nucleus	30	1.41	42.5	2.12	0.029	1.42
IPI00298961.3	XPO1 Exportin-1	XPO1	Cytoplasm	28.5	0.71	40	0.00	0.028	1.40
IPI00219446.5	PEBP1 Phosphatidylethanolamine-binding protein 1	PEBP1	Cytoplasm	18.5	0.71	23.5	0.71	0.019	1.27
IPI00005198.2	ILF2 Interleukin enhancer-binding factor 2	ILF2	Nucleus	58	0.00	49.5	0.71	0.037	0.85

IPI	Protein Description	Gene Symbol	Subcellular Localization	Average Spec-Count Control	STDev	Average Spec-Count Tamoxifen	STDev	ttest	Tamoxifen/ Control*
IPI00011913.1	HNRPA0 Heterogeneous nuclear ribonucleoprotein A0	HNRNPA0	Nucleus	45	1.41	38	1.41	0.038	0.84
IPI00021016.3	TSFM Isoform 1 of Elongation factor Ts, mitochondrial precursor	TSFM	Mitochondrion	23.5	0.71	19.5	0.71	0.030	0.83
IPI00027834.3	HNRPL heterogeneous nuclear ribonucleoprotein L isoform a	HNRNPL	Nucleus/Cytoplasm	85.5	3.54	67	2.83	0.032	0.78
IPI00009904.1	PDIA4 Protein disulfide-isomerase A4 precursor	PDIA4	Endoplasmic reticulum	78	1.41	60.5	2.12	0.016	0.78
IPI00303476.1	ATP5B ATP synthase subunit beta, mitochondrial precursor	ATP5B	Mitochondrion	262.5	6.36	202.5	3.54	0.017	0.77
IPI00171903.2	HNRPM Isoform 1 of Heterogeneous nuclear ribonucleoprotein M	HNRNPM	Nucleus	66	2.83	49.5	3.54	0.039	0.75
IPI00098902.4	OGDH oxoglutarate (alpha-ketoglutarate) dehydrogenase (lipoamide) isoform 1 precursor	OGDH	Mitochondrion	46	1.41	33.5	0.71	0.021	0.73
IPI00025086.3	COX5A Cytochrome c oxidase subunit 5A, mitochondrial precursor	COX5A	Mitochondrion	26.5	0.71	19	1.41	0.045	0.72
IPI00025057.1	ADAR Isoform 2 of Double-stranded RNA-specific adenosine deaminase	ADAR	Cytoplasm/Nucleus	28	1.41	20	1.41	0.030	0.71
IPI00027230.3	HSP90B1 Endoplasmin precursor	HSP90B1	Endoplasmic reticulum	154.5	4.95	109.5	0.71	0.045	0.71
IPI00100151.4	XRN2 Isoform 1 of 5'-3' exoribonuclease 2	XRN2	Nucleus	20.5	0.71	14.5	0.71	0.014	0.71
IPI00101186.6	RRP12 Isoform 1 of RRP12-like protein	RRP12	Nucleus	26.5	0.71	18	0.00	0.037	0.68
IPI00000877.1	HYOU1 Hypoxia up-regulated protein 1 precursor	HYOU1	Endoplasmic reticulum	81	2.83	55	1.41	0.020	0.68
IPI00329791.9	-	DDX46	Nucleus	28	1.41	19	1.41	0.024	0.68
IPI00006865.3	SEC22B Vesicle-trafficking protein SEC22b	SEC22B	Endoplasmic reticulum	18.5	0.71	12.5	0.71	0.014	0.68
IPI00215948.4	CTNNA1 Isoform 1 of Catenin alpha-1	CTNNA1	Cytoplasm/Cytoskeleton	58.5	2.12	39.5	0.71	0.032	0.68
IPI00027493.1	SLC3A2;LOC442497 4F2 cell-surface antigen heavy chain	SLC3A2; LOC442497	Apical cell membrane	47	1.41	31	0.00	0.040	0.66
IPI00024279.4	HEATR1 HEAT repeat-containing protein 1	HEATR1	Nucleus	24	1.41	15.5	0.71	0.037	0.65
IPI00069750.2	SIAHBP1 fuse-binding protein-interacting repressor isoform a	PUF60	Nucleus	31.5	2.12	20	1.41	0.033	0.63
IPI00291467.7	SLC25A6 ADP/ATP translocase 3	SLC25A6	Mitochondrion	24.5	0.71	15.5	0.71	0.006	0.63
IPI00216492.1	HNRPH3 Isoform 2 of Heterogeneous nuclear ribonucleoprotein H3	HNRNPH3	Nucleus	17.5	0.71	11	0.00	0.049	0.63
IPI00470467.5	POR NADPH--cytochrome P450 reductase	POR	Endoplasmic reticulum	25.5	2.12	16	1.41	0.045	0.63

IPI	Protein Description	Gene Symbol	Subcellular Localization	Average Spec-Count Control	STDev	Average Spec-Count Tamoxifen	STDev	ttest	Tamoxifen/Control*
IPI00305068.5	PRPF6 Pre-mRNA-processing factor 6	PRPF6	Nucleus	29.5	0.71	18.5	0.71	0.004	0.63
IPI00027252.6	PHB2 Prohibitin-2	PHB2	Mitochondrion/ Cytoplasm/Nucleus	52	0.00	32.5	2.12	0.049	0.63
IPI00240812.6	-	PDS5B	Nucleus	17	0.00	10.5	0.71	0.049	0.62
IPI00184330.5	MCM2 DNA replication licensing factor MCM2	MCM2	Nucleus	28.5	2.12	17.5	2.12	0.035	0.61
IPI00006196.3	NUMA1 Isoform 2 of Nuclear mitotic apparatus protein 1	NUMA1	Nucleus	86	1.41	51.5	4.95	0.048	0.60
IPI00003925.6	PDHB Isoform 1 of Pyruvate dehydrogenase E1 component subunit beta, mitochondrial precursor	PDHB	Mitochondrion	24.5	0.71	14.5	0.71	0.005	0.59
IPI00019269.3	WDR61 WD repeat protein 61	WDR61	-	9.5	0.71	5.5	0.71	0.030	0.58
IPI00867509.1	-	CORO1C	-	9.5	0.71	5.5	0.71	0.030	0.58
IPI00000155.6	-	N-PAC	Nucleus	11.5	0.71	6.5	0.71	0.019	0.57
IPI00021805.1	MGST1 Microsomal glutathione S-transferase 1	MGST1	Microsome	11.5	0.71	6.5	0.71	0.019	0.57
IPI00028387.3	C20orf116 Isoform 1 of Uncharacterized protein C20orf116 precursor	C20orf116	Secreted	11.5	0.71	6.5	0.71	0.019	0.57
IPI00032827.1	SF3B14 Pre-mRNA branch site protein p14	SF3B14	Nucleus	11.5	0.71	6.5	0.71	0.019	0.57
IPI00216654.2	NOLC1 Isoform Beta of Nucleolar phosphoprotein p130	NOLC1	Nucleus/Cytoplasm	36.5	3.54	20.5	2.12	0.048	0.56
IPI00013744.1	ITGA2 Integrin alpha-2 precursor	ITGA2	Membrane	34.5	0.71	19	1.41	0.016	0.55
IPI00021338.1	DLAT Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial precursor	DLAT	Mitochondrion	14.5	0.71	7.5	0.71	0.010	0.52
IPI00784154.1	HSPD1 60 kDa heat shock protein, mitochondrial precursor	HSPD1	Mitochondrion matrix	414	15.56	212	1.41	0.033	0.51
IPI00005158.1	LONP1 Lon protease homolog, mitochondrial precursor	LONP1	Mitochondrion	30.5	0.71	15.5	2.12	0.043	0.51
IPI00007188.5	SLC25A5 ADP/ATP translocase 2	SLC25A5	Mitochondrion	65.5	0.71	32	2.83	0.028	0.49
IPI00019848.2	-	HCFC1	Nucleus/Cytoplasm	20	1.41	9.5	2.12	0.038	0.48
IPI00029133.4	ATP5F1 ATP synthase B chain, mitochondrial precursor	ATP5F1	Mitochondrion	26	1.41	12	1.41	0.010	0.46
IPI00339384.5	RDH11 Isoform 1 of Retinol dehydrogenase 11	RDH11	Endoplasmic reticulum	15.5	0.71	7	1.41	0.037	0.45
IPI00019380.1	NCBP1 Nuclear cap-binding protein subunit 1	NCBP1	Nucleus/Cytoplasm	13.5	0.71	6	0.00	0.042	0.44

IPI	Protein Description	Gene Symbol	Subcellular Localization	Average Spec-Count Control	STDev	Average Spec-Count Tamoxifen	STDev	ttest	Tamoxifen/ Control*
IPI00095891.2	GNAS Isoform XLas-1 of Guanine nucleotide-binding protein G(s) subunit alpha isoforms XLas	GNAS	Cell membrane	11.5	0.71	5	0.00	0.049	0.43
IPI00383460.7	GRSF1 G-rich RNA sequence binding factor 1	GRSF1	Cytoplasm	22.5	2.12	9.5	2.12	0.026	0.42
IPI00023972.5	DDX47;APOLD1 Probable ATP-dependent RNA helicase DDX47	DDX47	Nucleus	15.5	0.71	6	0.00	0.033	0.39
IPI00219430.10	THRAP4 Thyroid hormone receptor-associated protein complex 100 kDa component	MED24	Nucleus	6.5	0.71	2.5	0.71	0.030	0.38
IPI00219483.1	SNRP70 Isoform 2 of U1 small nuclear ribonucleoprotein 70 kDa	SNRP70	Nucleus	9.5	0.71	3.5	0.71	0.014	0.37
IPI00215884.4	SFRS1 Isoform ASF-1 of Splicing factor, arginine/ serine-rich 1	SFRS1	Cytoplasm/Nucleus	25	0.00	9	1.41	0.040	0.36
IPI00337541.3	NNT NAD(P) transhydrogenase, mitochondrial precursor	NNT	Mitochondrion	14.5	2.12	5	1.41	0.045	0.34
IPI00337397.1	NUP98 Isoform 5 of Nuclear pore complex protein Nup98-Nup96 precursor	NUP98	Nucleus	7.5	0.71	2.5	0.71	0.019	0.33
IPI00009607.1	RAP2C Ras-related protein Rap-2c precursor	RAP2C	Cell membrane	8.5	0.71	2.5	0.71	0.014	0.29
IPI00152890.1	NOL6 Isoform 1 of Nucleolar protein 6	NOL6	Nucleus	8	0.00	1.5	0.71	0.049	0.19
IPI00063762.1	HPDL 4-hydroxyphenylpyruvate dioxygenase-like protein	HPDL	-	8.5	0.71	1.5	0.71	0.010	0.18
IPI00005614.6	SPTBN1 Isoform Long of Spectrin beta chain, brain 1	SPTBN1	Cytoplasm	15	1.41	2.5	0.71	0.021	0.17
IPI00216230.3	TMPO Lamina-associated polypeptide 2 isoform alpha	TMPO	Nucleus	18.5	0.71	3	0.00	0.021	0.16
IPI00465294.2	CDC5L Cell division cycle 5-like protein	CDC5L	Cytoplasm/Nucleus	11.5	0.71	1.5	0.71	0.005	0.13
IPI00465294.2	CDC5L Cell division cycle 5-like protein	CDC5L	Cytoplasm/Nucleus	11.5	0.71	1.5	0.71	0.005	0.13
IPI00412404.1	SUPV3L1 SUV3-like protein	SUPV3L1	Mitochondrion/Nucleus	6.5	0.71	0.5	0.71	0.014	0.08

*Ratio of average tamoxifen spectral counts over average spectral counts from the control sample.

Table 3

Spectral count comparison of DDA GPF vs. DIA PAcIFIC for a subset of identifications.

		GPF				PAcIFIC			
IPI	**Protein Description Control 1**	**Control 1**	**Control 2**	**Tamoxifen 1**	**Tamoxifen 2**	**Control 1**	**Control 2**	**Tamoxifen 1**	**Tamoxifen 2**
IPI00021439.1	ACTB Actin, cytoplasmic 1	150	142	144	163	218	215	183	163
IPI00304925.5	HSPA1A Heat shock 70 kDa protein 1	24	25	37	35	137	109	144	129
IPI00010951.2	EPPK1 Epiplakin	13	15	7	10	148	140	106	107
IPI00453476.2	hCG_2015269 similar to Phosphoglycerate mutase 1	7	8	11	12	23	30	34	27
IPI00847579.1	RPS12 ribosomal protein S12	6	6	4	8	15	16	16	12
IPI00329389.8	RPL6 60S ribosomal protein L6	2	4	1	5	32	24	29	23
IPI00007427.2	AGR2	0	0	1	1	2	0	11	12